

Fig. 1A

2701 TTATTATCCAGAACATCTGTTAATCATTACTCCAAACTAGACACTATTTACACACTCT
HNF1 HNF3

2761 ATGGAAGGCGGGTATATTATATAAGAGAGAAACACACATAGCGCCTCATTTGTGGTC
SpI TBP RNA Start

2821 ACCATATTCTTGGGAACAAAGATCTACAGCCATGGGGC
PreS1 protein start

Fig. 1B

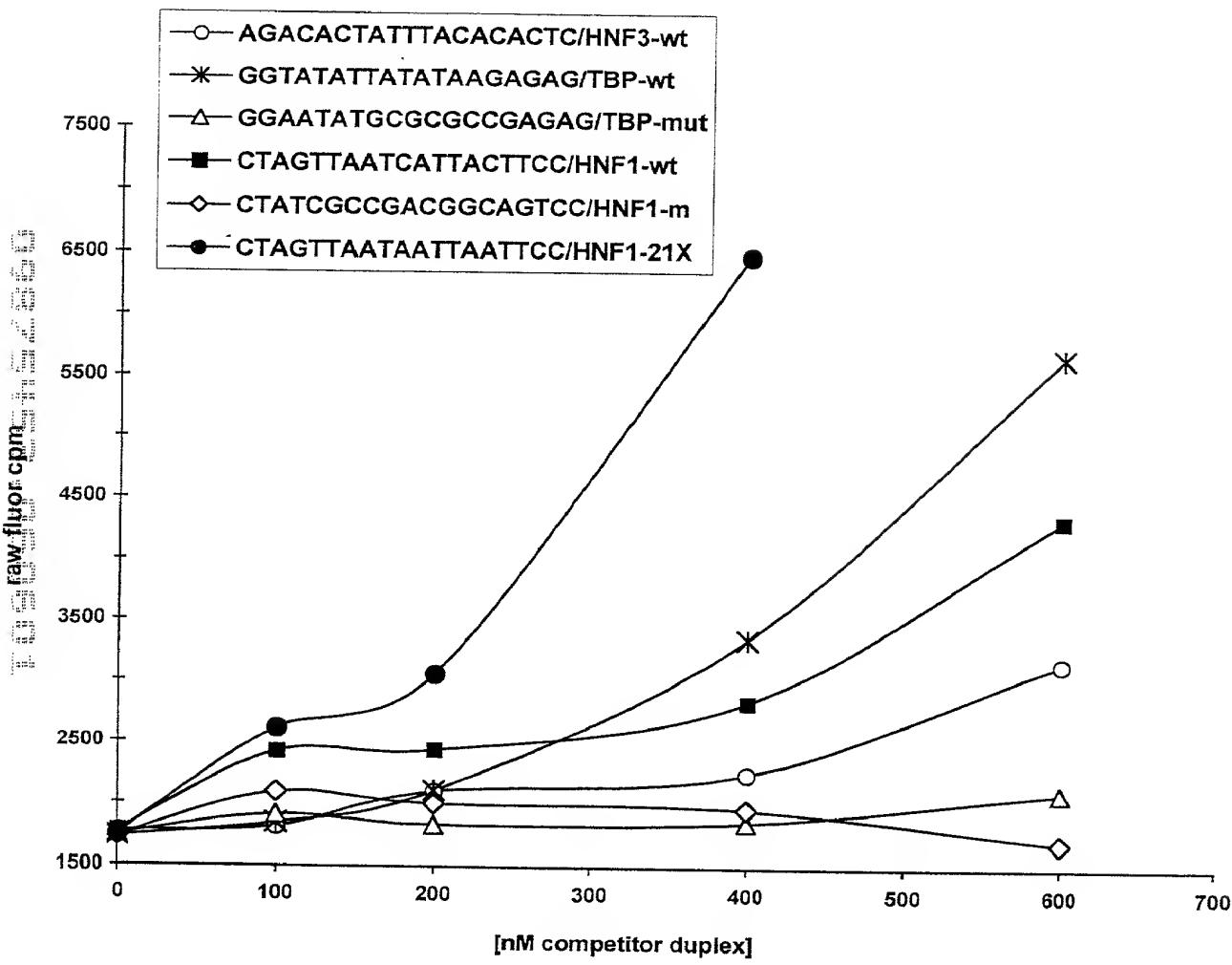


Fig. 2

1081 CTA AGC AGG CTT TCA CTT TCT CGC CAA CTT ACA AGG CCT TTC TGT GTA AAC AAT
NF1(1100-1119) 2c (1119-1134)

1135 ACC TGA ACC TTT ACC CCG TTG CCC GGC AAC GGC CAG GTC TGT GCC AAG TGT TTG
EF-C(1148-1168)

1189 CTG ACG CAA CCC CCA CTG GCT GGG GCT TGG TCA TGG GCC ATC AGC GCA TGC GTG
E (1180-1202) NF1(1209-1236) X-PBP(1229-1245)

1243 GAA CCT TTT CGG CTC CTC TGC CGA TCC ATA CTG CGG AAC TCC TAG CCG CTT GTT

1297 TTG CTC GCA GCA GGT CTG GAG CAA ACA TTA TCG GGA CTG ATA ACT CTG TTG TCC

1351 TAT CCC GCA AAT ATA CAT CGT TTC CAT GGC TGC TAG 1386

Fig. 3

CAGCTGGG CCGCCCTTGT GCGCGGGCTG ATGCTCTGAG GCTTGGCTAT
GC GG GGGGCCA ACGCGATTGT GGGTGCTCGG GGAGTGGGGG GGGGCACGAC CGTAGGTGCT
CCCTGCTGGG GCAACCCATC GCTCCCCATG CGGAATCCGG GGGTAATTAC CCCCCCAGGA
CCC GGAAATAT TAGTAATCCT AATTCCCAGC GGGGGAGGGG GCGCGGGAGG AATTCA CCT
GAAAGGTGGG GGTGGGGGGG GT CGCATCTT GCTGTGAGCA CCCTGGCGAA GGGGAGAGGG
CTTTTCTAT CAGTTTCTT TGAGCTTTA CTGTTAACAG GGTACGGTGG TTTGATGACA
CTGAACTATA TTCAAAAGGA AGTAAATGAA CAGTTTTCTT AATTTGGGC AGGTACTGTA
AAAATAAAA CAAAAGTTAA GACAGTAAAA TGTCCTTTA TTTTTAATG CACCAAAGAG
ACAGAACCTG TAATTTAAA AACTGTGTAT TTAATTTAC ATCTGTTAA GTT TGC GATA
ATATTGGGA CCCTCTCATG TAACCACGAA CACCTATCGA TTTTGTAAA AATCAGATCA
GTACACTCGT TTGTTTAATT GATAATTGTT CTGAATTATG CCGGCTCCTG CCAGCCCCCT
CACGCTCACG AATTCA GTCC CAGGGCAAAT TCTAAAGGTG AAGGGACGTC TACACCCCCA
ACAAAACCAA TTAGGAACCT CCGTGGTCTT GTCCCAGGCA GAGGGACTA ATATTCCAG
CAATTAAATT TCTTTTTAA TTAAAAAAA TGAGTCAGAA TGGAGATCAC TGTTTCTCAG
CTTTCCATT AGAGGTGTGT TTCTCCCGGT TAAATTGCCG GCACGGGAAG GGAGGGGGTG
CAGTTGGGA CCCCCGCAAG GACCGACTGG TCAAGGTAGG AAGGCAGCCC GAAGAGTCTC
CAGGCTAGAA GGACAAGATG AAGGAAATGC TG GCC ACCAT CTTGGCTGC TGCTGGAATT
TTCGGGCATT TATTTTATT TATTTTTGA GCGAGCGCAT GCTAAGCTGA AATCCCTTA
ACTTTAGGG TTACCCCTT GGGCATTG AACGACGCC CTGTGCGCCG GAATGAAACT
TGCACAGGGG TTGTGTGCC GGTCTCCCC GTCCTTGAT GCTAAATTAG TTCTTGCAAT
TTACACGTGT TAATGAAAAT GAAAGAAGAT GCAGTCGCTG AGATTCTTG GCCGTCTGTC
CGCCCGTGGG TGCCCTCGTG GCGTTCTTGG AAATGCCCG ATTCTGCCGG CTTGGATATG
GGGTGTCGCC GCGCCCCAGT CACCCCTCT CGTGGCTC CCAGGCTGC TGCTGCGCC
GCCTCCTAG TTGTCCCCTA CTG CAGAGCC ACCTCCACCT CACCCCTAA ATCCC GGGGG
ACCCACTCGA GGC GGACGGG GCCCCCTGCA CCCCTCTCC CTGGCGGGGA GAAAGGCTGC
AGCGGGCGA TTG CATTTC TATGAAAACC GGA ACTACAGG GGCAACTCCG CCGCAGGGCA
GGCGCGGC GTCAGGGATG GCTTTGGGC TCTGCCCCTC GCTGCTCCG GCGTTGGCG
CCCGCGCCCC CTCCCCCTGC GCCC GCCCCCCC GCCCCCTCC CGCTCCATT CTCTGCCGGG
CTTGATCTT TGCTTAACAA CAGTAACGTC ACACGGACTA CAGGGAGTT TTGTTGAAGT
TGCAAAGTCC TGGAGCCTCC AGAGGGCTGT CGCGCAGTA GCAGCGAGCA GCAGAGTCCG
CACGCTCCGG CGAGGGCAG AAGAGCGCGA GGGAGCGCGG GGCAGCAGAA GCGAGAGCCG
AGCGCGGACC CAGCCAGGAC CCACAGCCCT CCCCAGCTGC CCAGGAAGAG CCCCCA

Fig. 4

10	20	30	40	50	60	70
GAATTCACTG	GGGAGAGCAT	TCAGGAAGAT	GACAACAGGA	TAATAGGTCA	ACAGAGTAAT	AGAGAGGTCG
CTTAAGTGAC	CCCTCTCGTA	AGTCCTTCTA	CTGTTGTCCT	ATTATCCAGT	TGTCTCATTA	TCTCTCCAGC
80	90	100	110	120	130	140
CTAAAAATAA	ACTCTAAGAA	GTATTCAAGCC	AAAACATTAA	TTGAGCTAAT	AATGGTGGGA	TCAATTTCAG
GATTTTTATT	TGAGATTCTT	CATAAGTCGG	TTTGATAAT	AACTCGATTA	TTACCACCC	AGTTAAAGTC
150	160	170	180	190	200	210
GGGAATATTG	TGGGCAGAAG	TCAGACTGTA	GGAGGCTGGG	GATCAAGAAG	TTGAGGCAAG	GAGGTTGGAC
CCCTTATAAAC	ACCOGTCTTC	AGTCTGACAT	CCTCCGACCC	CTAGTTCTTC	AACTCCGTT	CTCCAACCTG
220	230	240	250	260	270	280
AACAACGTGTT	TTTCAAGTT	GGTCACGTGA	ACAAATCTGT	GACCTTCAGC	CTCCCCTCCC	TCGGGTCTTG
TTGTTGACAA	AAAAGTTCAA	CCAGTGCACT	TGTTTAGACA	CTGGAAGTCG	GAGGGGAGGG	AGCCCAGAAC
290	300	310	320	330	340	350
GCTGAGCTGA	TTGCAGGGCC	CCTGCAGCTC	TGGCACTCTC	AAGTTGTATA	AAACTGACAG	TGCAGAAAGTC
CGACTCGACT	AACGTCCCAG	GGACGTCGAG	ACCGTGAGAG	TTCAACATAT	TTTGACTGTC	ACGTCTTCAG
360	370	380	390	400	410	420
CTTGAGGCCA	TTTTGGCTCT	CATGATAATT	TTCCTTCAGT	GGAACATAAGG	TTACTTGTCT	AAGAACCAAA
GAACTCGGGT	AAAACCGAGA	GTACTATTAA	AAGGAAGTC	CCTTGATTCC	AATGAACAGA	TTCTTGGTTT
430	440	450	460	470	480	490
GCCTCTGACT	TGACTGATCA	AAGTTCATCA	CGTGCATCGA	AGCCACCTAC	TTGGCAGATG	TAGTGAAAAG
CGGAGACTGA	ACTGACTAGT	TTCAAGTAGT	GCACGTAGCT	TCGGTGGATG	AACCGTCTAC	ATCACTTTTC
500	510	520	530	540	550	560
CTACATAGAT	CTGGGCCAG	GACAGGATGC	TGGGGCGTGG	GAGGGGAAGA	AAGCAGGTGC	TAACTATATA
GATGTATCTA	GACCCGGGTC	CTGTCCTACG	ACCCCGCACC	CTCCCCTTCT	TTCGTCCACG	ATTGATATAT
570	580	590	600	610	620	630
GATAGCATGC	CTATCAGAGC	AGTTTTACG	TTTCCTATT	GTCTCTCAAA	ACAATTTTAT	AGGAATCATC
CTATCGTACG	GATAGTCTCG	TCAAAAATGC	AAAGGATAAA	CAGAGAGTTT	TGTTAAAATA	TCCTTAGTAG
640	650	660	670	680	690	700
AAAGCAATT	TATCATGGTT	TCTAGACCAG	GTGGATGT	GAGGTAGGG	TTTCCACAGC	TGCTTTAGT
TTTCGTTAAA	ATAGTACCAA	AGATCTGGTC	CAAACCTACA	CTCCATCCCT	AAAGGTGTCG	ACGAAAATCA
710	720	730	740	750	760	770
TTGAAGGAAA	TCTGATAAGA	TGATGAAAAA	GCCCTTCAGA	AATGTGTAAT	CCTACACACT	TCAGTGATTC
AACTTCCTT	AGACTATTCT	ACTACGTTT	CGGGAAGTCT	TTACACATTA	GGATGTGTGA	AGTCACTAAG
780	790	800	810	820	830	840
AATTCAATTG	CAAAACTAA	GGTGTTTTA	ATATTGTTAT	TGTTCATTTG	GTGTTTACCA	ACATGTAAGG
TTAAGTAACA	GTGGATT	CCACAAAAT	TATAACAATA	ACAAGTAAAC	CAAAATGGT	TGTACATTCC
850	860	870	880	890	900	910
AGTTGGCAAT	TATTTGTTAA	ACTCATGTCT	TAGGCTAAAT	AAATTCCAAA	AAATTCAAGG	TGAGAATTGT
TCAACCGTTA	ATAAACAAATT	TGAGTACAGA	ATCCGATT	TTAAGGTTT	TTAAGTCCT	ACTCTTAACA

Fig. 5A

920 930 940 950 960 970 980
 TTATTGCTTA ACGTGTGTCA AATTTCTTC ATGCACATCT TTATTAGATC TTCACAGCAA CCTACAGGAT
 AATAACGAAT TGCACACAGT TAAAGAAGG TACGTGTAGA AATAATCTAG AAGTGTGCGTT GGATGTCCTA

 990 1000 1010 1020 1030 1040 1050
 AAGCAAGACA GGTGCAAGTG CCTCCTTTGG GTATGAGGAA ACTGAGGTCT AAAGAGATGA AGTGATTTGC
 TTCGTTCTGT CCACGTTCAC GGAGGAAACC CATACTCCTT TGACTCCAGA TTTCTCTACT TCACTAAACG

 1060 1070 1080 1090 1100 1110 1120
 CCAAGGCTCA TAGCAATTAA TTGGTAGAGC AAAGACTAGA ATTCTCTTAA CTGCAGCCTA TTTTCCCTAT
 GGTTCCGAGT ATCGTTAAAT AACCATCTCG TTTCTGATCT TAAGAGAATT GACGTCGGAT AAAAGGGATA

 1130 1140 1150 1160 1170 1180 1190
 TCTGAACTGT TACATCAGCA TCAACAAATTAA TCTAATGGAT TGGAACAGTG TACACAGGCA GCTTAGCTAC
 AGACTTGACA ATGTAGTCGT AGTTGTTAAAT AGATTACCTA ACCTTGTAC ACATGTCGGT CGAATCGATG

 1200 1210 1220 1230 1240 1250 1260
 GTCAAGTCAC GATTTTACT TAAACTCAA TTCCAGAGTC TTGGCCTGAT TTCCCTCAAG ACCCTACTTA
 CAGTTCACTG CTAAAATGA AATTGAAGTT AAGGTCTAG AACCGGACTA AAGGGAGTTC TGGGATGAAT

 1270 1280 1290 1300 1310 1320 1330
 TCTTTGGCTT TGGAAAATTAA TTTTTCTTG CATTATCTTT CCAGCTAAAT TTTATTTAAAT AACCATCAGC
 AGAAACCGAA ACCTTTAAA TAAAAAGAAC GTAATAGAAA GGTCGATTAA AAATAAAATTA TTGGTAGTCG

 1340 1350 1360 1370 1380 1390 1400
 ATGCTTTTTTG TGCTTATGC CATGTAGACT TGACCTGAAA ACCTGCCAGG CTTTCATTGA GTTTAGTGT
 TACGAAAAAAA ACGAAATACG GTACATCTGA ACTGGACTTT TGGACGGTCC GAAAGTAACT CAAATCACTA

 1410 1420 1430 1440 1450 1460 1470
 TAAAGAAGTA AAGTTCTGAG AAGCAATTAG TTGATGGGAC ACCAGTCATA AAATCAATCC AAACCTTTGT
 ATTTCTTCAT TTCAAGACTC TTCGTTAAC TAACTACCCTG TGGTCAGTAT TTTAGTTAGG TTGAAAACA

 1480 1490 1500 1510 1520 1530 1540
 TGACATGTGT TTCTTCTCC ATATACCAGG TTCCCGCTTC GTATTAGTAA GATTGAAATT GAAATAAGTC
 ACTGTACACA AAGAAAGAGG TATATGGTCC AAGGGCGAAG CATAATCATT CTAACTTAA CTTTATTCA

 1550 1560 1570 1580 1590 1600 1610
 TATTGCTGGT GGATGAATTG GTCACTTTCC TTGAAACTGG TGAACCCAAA AAGTTAGACA GTGATAGGAA
 ATAACGACCA CCTACTTAAAC CAGTGAAGG AACCTTGACC ACTTGGGTTT TTCAATCTGT CACTATCCT

 1620 1630 1640 1650 1660 1670 1680
 AATACTGCCA TTGCTGTAA AGAACTCTAT GACATTCAA GGCAAGAATG AATATATGGA AGAAGAAACT
 TTATGACGGT AACAGACAAT TCTTCAGATA CTGTAAAGTT CCGTTCTAC TTATATACCT TCTTCTTTGA

 1690 1700 1710 1720 1730 1740 1750
 TGTTCCTTCT TTACTTACAA AAAGGAAAGC CTGGAACTGAA ATGATATGGG TATAATTAAA AAAAAAA
 ACAAAAGAAGA AATGAATGTT TTTCCCTTCG GACCTTCACT TACTATACCC ATATTAATTT TTTTTTTT

 1760 1770 1780 1790 1800 1810 1820
 AAAACAAAAAA ACCTTACGT AACGTTTTGC TGGGAGAGAA GACTACGAAG CACATTTCC AGGAAGTGTG
 TTTTGTGTTT TGGAAATGCA TTGCAAAACG ACCCTCTCTT CTGATGCTTC GTGTAAAAGG TCCTTCACAC

Fig. 5B

1830	1840	1850	1860	1870	1880	1890
GGCTGCAACG	ATTGTGCGCT	CTTAACATAAT	CCTGAGTAAG	GTGGCCACTT	TGACAGTCTT	CTCATGCTGC
CCGACGTTGC	TAACACGCGA	GAATTGATTA	GGACTCATTC	CACCGGTGAA	ACTGTCAGAA	GAGTACGACG
1900	1910	1920	1930	1940	1950	1960
CTCTGCCACC	TTCTCTGCCA	GAAGATACCA	TTTCAACTTT	AACACAGCAT	GATCGAAACA	TACAACCAAA
GAGACGGTGG	AAGAGACGGT	CTTCTATGGT	AAAGTTGAAA	TTGTGTCGTA	CTAGCTTTGT	ATGTTGGTTT
1970	1980	1990	2000	2010	2020	2030
CTTCTCCCCG	ATCTGCGGCC	ACTGGACTGC	CCATCAGCAT	GAAAATTTT	ATGTATTTAC	TTACTGTTT
GAAGAGGGGC	TAGACGCCGG	TGACCTGACG	GGTAGTCGTA	CTTTAAAAAA	TACATAAAATG	AATGACAAAA
2040	2050	2060	2070	2080	2090	2100
TCTTATCACCC	CAGATGATTG	GGTCAGCACT	TTTGCTGTG	TATCTTCATA	GAAGGCTGGA	CAAGGTAAGA
AGAATAGTGG	GTCTACTAAC	CCAGTCGTGA	AAAACGACAC	ATAGAAGTAT	CTTCCGACCT	GTTCCATTCT
2110	2120	2130	2140	2150	2160	2170
TGAACCACAA	GCCTTATTAA	ACTAAATTG	GGGTCTTAC	TAATTCTAG	GTTGGTTCTA	CCCAAATGAT
ACTTGGTGT	CGGAAATAAT	TGATTAAAC	CCCAGGAATG	ATTAAGTATC	CAACCAAGAT	GGGTTTACTA
2180	2190	2200	2210	2220	2230	2240
GGATGATGGT	AGAAACCAAA	TAGAAGAATG	GTCTTGTTG	ATAATGTTT	TTCCCTAGTC	AATGAACTCT
CCTACTACCA	TCTTTGGTTT	ATCTCTTAC	CAGAACACCG	TATTACAAAC	AAGGGATCAG	TTACTTGAGA
2250	2260	2270	2280	2290	2300	2310
CATATTCTTG	TCTCTGGTTA	GGATCTGGG	ATCTGGAGTC	AGACTGCCTG	GGCTCAAATC	TTGGCTCTGC
GTATAAGAAC	AGAGACCAAT	CCTAGAACCC	TAGACCTCAG	TCTGACGGAC	CCGAGTTAG	AACCGAGACG
2320	2330	2340	2350	2360	2370	2380
CCATACCATC	TCTGTTATCC	TGGGGCAAGT	GCCTCAGTT	CCACATCTGA	GAAATGGGGA	TGGTAGTGGT
GGTATGGTAG	AGACAATAGG	ACCCCGTTCA	CGGAGTCAAA	GGTGTAGACT	CTTTACCCCT	ACCATCACCCA
2390						
GTCCATTCA	TAGAT					
CAGGTAAAGT	ATCTA					

Fig. 5C

GAGATGTATAATAATTAGAAAAATCTCAAGGTTATCTTACTTTCTTA
GGAAATTAACAATTAAATATTAAGAAACGGCTCGTTCTTACACGGTAGACTTA
ATACCGTAAGAACGAGCCGTTTCGTTCTCAGAGAAAGATTGACAAGATTAA
CCATTGGCATCCCCGTTTATTGGTGCCCTTCACAGAAAGGGTTGGTCTTAA
TT

Fig. 6

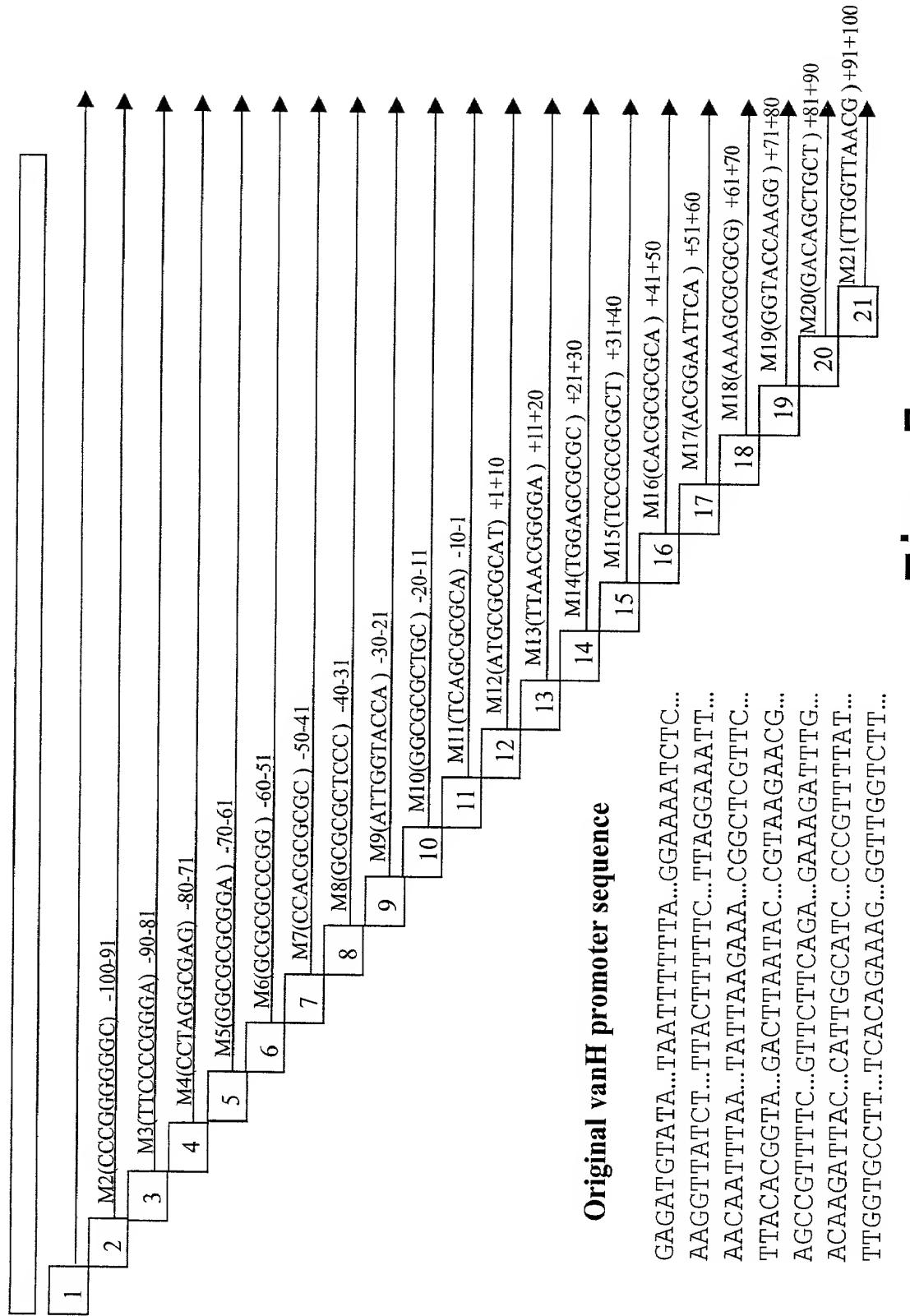


Fig. 7

TCTAGAAAAT AATTCCCAAT ATTGAATCCC AAAGAATTCA ACATTTGGGC TGTCGTTGA 61
 AAGATAAGTT GAATTGGTC ATGAAGGAAG AGAGGGGGGA TACAATTCA GTAAAAGGTA 121
 ACAGCAAGGT CCAAAGACAG TCAGGTCTTC AGTAGTATGG AGTATATTCA GAGGGAGCCA 181
 AGATGTCTGA TGTGAACTAA AAAGATTGGT GGTTGGTAGG AGGAAGAGGT GTGAGAAGAG 241
 GCTGTAAAGA AAAATTGAAA CTTGATTGTG ATGGACTTTA AAGGCTAGGC TATGGACTT 301
 GGACATGAAT CTGCAGGCCA GTGTTTCAG ACTGGCCCCC ATAACGTCT ATCACAGCAA 361
 CACAGACATG TGTTGTTGG CCTGCAGAGG TTGGCCTGC ATGATGATT TAAACCATCT 421
 GAATTAGTAG CCATCATTT CAAAAATCAA GAGATGCCAC ATTAATAAT GGAATGCTGC 481
 TGTCTTGA AATAATGAAA CATCTGGAAC ATTGAGGCCA CATTCCTGAC TGACAGCAAT 541
 CAGTTGGAGC TGCAGTAGTGA CTGCCACTT TACATGGGGC ATCTGATCCC TAGTCGATTA 601
 CAGCTGCCAC CACTCCCTT TATCTCTCTA ATACCAAGCT CTTTCACTC ATTTTTGTTA 661
 CTTAAGAGAT ATTTGGGTTT GAAACCTCTG ATGCAGGTAA TTGAGGGTTA TAGAGCAGAG 721
 GACAGATGCT ATCAGAGTTG TCTTTAAGA AAGAACCCCTC TGTTCTTCAT TTTGTTGAAG 781
 ATAGCCTGGA AGAGGGCAGC CAGGGGAGAA GTTAGGGCTG GAGCTATGAG AAAGCATAAG 841
 ATGAGATGAT GGCTTCAACA TTGAGGACAG AAAGAATATT GAGATGAGAA AGTAGTCCAT 901
 ATAAGCATCT ATGCAAAGGA AATAGCAGAT GTCCTCAAAT CAGCAGAGGC AACAACTCTG 961
 AAAGTTTATT CATAAGCCCC TCTTTCATC TCCAATCCAG TTCAAATGTA ATTATTAAA 1021
 TTGTTCTTCAT CTCTCCTTCC TGGATCATGA ATGAGCTCCT TAAATGCAGG GTCCACAGTG 1081
 TCCTATTCTAT CAGTGAATTCAAGTGCCTA GCACAGAGCC TGGCAAATAG TAAATGCTTA 1141
 ACAAAATATTC GTTCAGTGCA TGAATTGGAG TGATTCTCTA CTTGCTCTCA TAAGTTGAAA 1201
 AAAGGTTTAT TACATACCTA AATATGCTGA AATCACAGGG CATTGGCAA CCCCCCAAAA 1261
 CCAAAACTCC CAGTTTGGAA ACAGAATTTT AATTCTGTGA AAATAAAATC CATTCTTTA 1321
 TTCAAAAAAT ATTTATTAAA CAATGACCAT GTCCACACCA GGCTGAGTCC TAAGGATTCA 1381
 ATGATGAACA AAAACCAACA TGATTCTGC TCTTAGGAAA CATAACAGTTC AGTGAGGAAA 1441
 ACAGATTGTG AGAAGTCCTC CAACAAATAC TGGGTGCTAT TAAAATATAT TAAAAGGTGA 1501
 GTGGGTGAGG GACTTGAGCT AGCCTAGGTG GTTCAGGAAG TCTCCTGGA TGTGCTGATA 1561
 TGCATAGGCA TTAACTAGAT AAATAGAGAG AAGGATGAAC CAACATTGCA GGTAGAGGGA 1621
 ACAGAATATG CAAAGGCAGG AAGGATTATG GAGTCGTTGG AGGACCTGAA TAAAGGCCA 1681
 GTGTAAGTGG ATCTCAGAAA ACAGGAGGAA AGGTGTATGA GATGAGATCA GAGAGGCAGA 1741
 TCATGTGGGG TATGGTTAAT GTTTTGGACT TTTCTATTAA GAGCAATGGG GAGACAGTGA 1801
 CAGGACTTAA ACGGGGAAAT AATATGACCA GATTAAACTT TCTAAAAAAC CCTCTATGCA 1861
 AATATATATT GAGAGTTAAT TATTGACAAA GATTCAAAGG CAACAAAGTG GAGAGAGAAT 1921
 AGTATTTCA AAAAATGGTG CCAAAACAAT AGGACATCTA TATTAAGT TGGGTATCTG 1981
 TCTACAAAAC TTAATTCAAA ATGGATCACA GACCTAAATG TAAAACGTAA AGCTATACAA 2041
 CTTCTGGAAG GAAAACACAG ATGGAATCT GTGTGATCTT GAGTTGAAA ATGATTATT 2101
 ATATCTGACA CCATAATCCG TAAGTTAAC TAATTCTAA GTGAACAAAG TGATGAACTG 2161
 GACTCATCA GAATTAAAA TGTTGTGCT TCAAAAGACA CTGGTATGAT AATGAAGACA 2221
 AACTACAGAT AAGATATTGT TGAATCATAT TTCTGATAAA GGAATTGTGG CTCAGAATAC 2281
 ATAACCTCTAA ACCCCCCATAA TAAATTACAA GTAGCCCAAT TAAAAAAAAA AAAAGAGAAA 2341
 AAATTACAG TCTTCATCAA AGAAAGTATC AATTGTAAAA TAAGCACATG AAAATGCTC 2401
 TGCATCTTAA TTCATGGGG GATGAAATAA AAATTAAATG GGAAAGACAC CTCTAATTAG 2461
 AATACTAAAA TTAAAAAAGAC TGACCATACC AAGTATTGGT GAAGTGGAAA TGTAAAATGA 2521
 TACAATCAAC TTAGGTAGAT GATTGGAAG TTTCTTACAA AAGTAGGTGT ATACCTACCC 2581
 TGTGACTCAC CCATTCCATG GCTAAGTATT TACCTGAGAG AAATGAAAGA ATACATCCAT 2641
 ACAAAAGATGT TTATACAAAT ATTATAGCA GTTTTATTG TAGTAGCCCC AACTGAAAAA 2701
 GAACCCAAAT GTCCATCAA AGTGAATGGA TAAACAAAGC GTGGTACAGC AATGCAATAG 2761
 AATACTACTT AGCAATAAAAG AAGAATGAGC TAGTGTATATA CATAACAGCT TAAATGTACA 2821
 TCAAAGGCAT TGTGCTCACT GAAAGATGCA AGTAAAAAAA AAAAGAGTA CATGCTGTAT 2881
 AGTTCATTG ACATAAAAAT CTGGAAAGTG AAAAACAGTC TATACTGACA GAAAGCAGAT 2941
 CATTGGTTGC CTGAGGAGGA GGAGTATAGG AGAGGTGGAG GGAAAATGTA CAAAGTGGCA 3001
 CAATAAAAAC TTTTGGAAATC ATAGATATAT TCACTATCTT GATTGAGTGA TGATTTCATG 3061

Fig. 8A

AGTGCACGTG CGTGTGTCAA AAATGATCAA TTTATGCAAC TTTAAATATG TGCAGTTAT 3121
 TGTATATATC AATTATACT CAGTACGGCT ATTAAAAAGA AACCCCTCTGG CTGCACAATG 3181
 CAGAACTGAT TCTAGGAAAG AGTGGAGGGA GGATGACCAT TTACAGTGCT CCAGGTGGAA 3241
 GAGAACGGTG CCTCTGGAA GTGAACTAGG TTGGCAACAA CAGAGATGAA ATAAATGGC 3301
 AGATGTGTGA GATACTTAGG AAATAAAACC CGATGGTCAC CATTTCCAA AGGTCAAGCTC 3361
 ATCCCTGGCTT TCCAGAGCAA AGAGCTAGGG AAGACTTTAT TAATAAATCC CTCTTGAAAGT 3421
 TGCGAGGAA GCTTATAGCA GAAACTTACT CTCAACCTGA CTAATCTGAG AGAACACCTC 3481
 TGGTTCCATT TGATTACTAA AAAACTGCAA AGAACAGGAG GAGAAAGAAG AAGAAAGCTG 3541
 GTACAAACAG TGAACCTATA TAATATTAAT CAATAATTGT CTCTTGTTCT TAAAAGCAAT 3601
 GGGAAAGAAAA TGAGATTGAG GCTGGAAGAT CAGAGTCAA AATCCAAATA AAGTATATGG 3661
 CCCTAATATG CTTATAGTAG TTAACCTTTC CTGATAATGA TATAATTGTT GACAGCACCA 3721
 TCTTTAAAT AAAATAACAT AGTAATCCTT CAGATTGTA GAAGATCTT CCTGTTACA 3781
 AGTTTGTCT ATACACATTA TGTCTTTAA ATGACACACT AGCCTCTGA GGGTAACCTA 3841
 TATTGGCAAC AGTTTCAGA TGTGGAAACT GTGAAGACAA TGTGTTGTAT GTGGAAGCAA 3901
 CATAAAACTT GGAGTCTTC AGACCCAGGT TTGAATGTCA GACTGCTTT TATTCAAGT 3961
 AACTTCAGAG CATTATTCTT CACCTTAATT TTTTTTCAGG CCTCTTTGTG TCTATGTGTC 4021
 CTCTTCACTC CTGTCATTG TTTCTTCAGT GATTTTGCC ACCTTCCTTC ACTGTTAGTG 4081
 TGTAGACACA TAGTTCTCCT GGCTCTGAGA GCCTATGTTA ATTCCATTCT ACCATCCTGC 4141
 CACGGCCAC TCAATTCTA TTGAGCAATG CTAGTTGAAA GTTGTGGTGG GATTAAATGT 4201
 TGCAATGAGT ATTCAAATGA GGTGAAAGTA TCTACGCATT CTACTTACAT ATGGTGAGGT 4261
 ATATTCAAGG AAGCTGTAGC CATAAAATC TCAGGAAATA ATTTTCACC TCCTCAGGTG 4321
 AAAGGGTCTT CAGGCCTTG TGTCTGGAA GTTTCATTAA TAGCCATTTC CCAAATGACA 4381
 ATGCGATTGA TGAGTCTAGA GTCTAGCTCA AATAGCAATG GACTGGAAGA CTAGTTAGG 4441
 TTTTACTAAT GTGGAACATA GAACAAATTAA TGTCTTGTT TCAGCCTGTT CATCTGTGAA 4501
 ATAGAGCCTA TCATATCCAG TCTCCTTGC CTTTAGGTTT GAGTTACCTT CTTGGTCAA 4561
 GGTAAGTAA TGCTATGAT GTTGGCTGT GCACAAGATA AAGCTACAAC AAAGCTACAA 4621
 CCCATCTTT CTCTGTAGAA GACTCAAAAA GCAAAAGAGA CCCAGGAAA TCTCGGAATG 4681
 ACTTTGGAA CAGAGGCCCT CCCCAGAACAGAAGTCAAAG GAATTAAAC ATAGGGAAGG 4741
 CCCAGGTCTC TACTGACATA AAGGAAAGAT GTTTCTTAT AGGTTACAGT TTTACATTT 4801
 CTCTCTCTTG ATCCCATTCC CACTTGCATC TGCCACCTTT ACACAGGGCT TATGGGACCT 4861
 CCTCCACAA AGAGCAGTT CAGTAACCCA CATCATCCTC TACGCCCTGG CTGTCCATCA 4921
 AGAGGCAGAA AGCAGCCCTA TATAGTTCT ATCCTGGAT AGTCCAGTT GTAAAGTTA 4981
 AAATATGCGA AGGCAACTG GAAAAGCAAG CGGCTGCATA CAAAGCAAAC GTTACAGAG 5041
 CTCTGGACAA AATTGAGCGC CTATGTGTAC ATGGCAAGTG TTTTAGTGT TTGTGTGTT 5101
 ACCTGTTGT CTGGGTGATT TTGCTTTGA GAGTCTGGAG AGTAAAGTA CTGGTTAAAG 5161
 GAACTTCCAG ACAGGAAGAA GGCAGAGAAG AGGGTAGAAA TGACTCTGAT TCTTGGGCT 5221
 GAGGGTTCT AGAGCAAATG GCACAATGCC ACGAGGCCCG ATCTATCCCT ATGACGGAAT 5281
 CTAAGGTTTC AGCAAGTATC TGCTGGCTTG GTCATGGCTT GCTCCTCAGT TTGTAGGAGA 5341
 CTCTCCCCTC CTCCCATTG CGCGCTCTTA TCAGTCTGA AAAGAACCCC TGGCAGCCAG 5401
 GAGCAGGTAT TCCTATCGTC CTTTCCCTCC CTCCCTCGCC CCACCCCTGTT GTTTTTTAG 5461
 ATTGGGCTTT GGAACCAAAT TTCTGAGTG CTGGCCTCCA GGAAATCTGG AGCCCTGGCG 5521
 CCTAAACCTT GGTTTAGGAA ACCAGGAGCT ATTCAAGGAAG CAGGGGCTCT CCAGGGCTAG 5581
 AGCTAGCCTC TCCTGCCCTC GCCCACGCTG CGCCAGCACT TGTTCTCCA AAGCCACTAG 5641
 GCAGGCCTTA GCGCGCGGTG AGGGGAGGGG AGAAAAGGAA AGGGGAGGGG AGGGAAAAGG 5701
 AGGTGGGAAG GCAAGGAGGC CGGCCCGGTG GGGCGGGAC CGCACTCGCA AACTGTTGCA 5761
 TTTGCTCTCC ACCTCCCAGC GCCCCCTCCG AGATCCCAGG GAGCCAGCTT GCTGGGAGAG 5821
 CGGGACGGTC CGGAGCAAGC CCACAGGCGAG AGGAGGGAGC AGAGGGAAAA AGGGCCGAGC 5881
 TAGCCGCTCC AGTGCTGTAC AGGAGCCGA GGGACGCACC ACGCCAGCCC CAGCCCGGCT 5941
 CCAGCGACAG CCAACGCCTC TTGAGCGCG GCGGCTTCGA AGCCGCCGCC CGGAGCTGCC 6001
 CTTTCCCTTT CGGTGAAGTT TTTAAAGACT GCTAAAGACT CGGAGGAAGC AAGGAAAGTG 6061

Fig. 8B

CCTGGTAGGA CTGACGGCTG CCTTTGTCCT CCTCCTCTCC ACCCCGCCTC CCCCCACCCCT 6121
GCCTTCCCCC CCTCCCCCGT CTCTCTCCC GCAGCTGCCT CAGTCGGCTA CTCTCAGCCA 6181
ACCCCCCTCA CCACCCCTCT CCCCACCCGC CCCCCCGCCC CCGTCGCCCA GCGCTGCCAG 6241
CCCGAGTTG CAGAGAGGTA ACTCCCTTG GCTGCGAGCG GGCGAGCTAG CTGCACATTG 6301
CAAAGAAGGC TCTTAGGAGC CAGGCAGCTG GGGAGCGGCT TCAGCACTGC AGCCACGACC 6361
CGCCTGGTTA GGCTGCACGC GGAGAGAACCTCTGTTC CCCCCACTCTC TCTCCACCTC 6421
CTCCTGCCTT CCCCCACCCCG AGTGCAGGAGC CAGAGATCAA AAGATGAAAA GGCAGTCAGG 6481
TCTTCAGTAG CCAAAAAACA AAACAAACAA AAACAAAAAA CAAGAAATAA AAGAAAAAGA 6541
TAATAACTCA GTTCTTATT GCACCTACTT CAGTGGACAC TGAATTGGA AGGTGGAGGA 6601
TTTGTTTTT TTCTTTAAG ATCTGGGCAT CTTTGAATC TACCCCTCAA GTATTAAGAG 6661
ACAGACTGTG AGCCTAGCAG GGCAGATCTT GTCCACCGTG TGTCTCTTC TGCACGAGAC 6721
TTTGAGGCTG TCAGAGCGCT TTTGCGTGG TTGCTCCCGC AAGTTCTT CTCTGGAGCT 6781
TCCCGCAGGT GGGCAGCTAG CTGCAGCGAC TACCGCATCA TCACAGCCTG TTGAACCTT 6841
CTGAGCAAGA GAAGGGGAGG CGGGGTAAGG GAAGTAGGTG GAAGATTCAAG CCAAGCTCAA 6901
GGATG

Fig. 8C

CA GGCCCCACAA AACCTAGATC TGCCCCAGTA TAACTAAATC 1501
 TGGGACCATT TATTGAGCAA TTATTATGTG CCAAGTATTG CGCTGAGTGC TTCCAGAGCA 1561
 TTATCTCCTT TAACCCCAGC ATAGTATGTC AGATGCTGTT TTACAGATGA GCCAACTGAG 1621
 ACCAGAGATG CTCAGTCACT TGCCAAGGT GACATGACTG ATATGGAATA GAGTCAAGAT 1681
 TTTTTTTTTT TTTTTGACA CGGAGTCTCA CTCTGCTCTC CAGGCTGGAG TGCAGAGGCG 1741
 CAATCTCAGC TCACTGCAAG CTCTGCCCTCC CAGGTTCACG CATTCTCCTG CCTCAGCCTC 1801
 CTGAGTAGCT GGGACTACAG GCACCCGCCA CCACACCTGG CTAATTTTT GTATTTTAG 1861
 CAGAGACAGG GTTTCACCCT GTTAGCCAGG ATGGTCTCGA TCTCCTGACC TCGTGATCTG 1921
 CCTGCCTCGG CCTCCCAAAG TGATGGAATT ACAGGTGTGA GCCACCGCGA CTGGCCAGAT 1981
 TCAAGATTG AACCCAGGTC CTCTGGTCC CAGAGGCCCG TGTTTCTCAA CTCCCTAGCA 2041
 TGCATACGCA CCTGTCCCTC TAGAGGTGCC TGCTTAAGTG TGCTCAGCAC ATGGAAGCAA 2101
 GTTAGAAATG CTAGGTATAC CTGTAAAGAG GTGTGGGAGA TGGGGGGGAG GGAAGAGAGA 2161
 AAGAGATGCT GGTGTCCCTC ATTCTCCAGT CCCTGATAGG TGCCTTTGAT CCCTTCTG 2221
 CCAGTATAGC TGCATTCTTG GCTGGGGCAT TCCAACCTAGA ACTGCCAAAT TTAGCACATA 2281
 AAAATAAGGA GGCCCAGTTA AATTGAAATT TCAGATAAAC AATGAATAAT TTGTTAGTAT 2341
 AAAATATGTCC CATGCAATAT CTTGTTGAAA TTAAAAAAA AAAAAAAAGT CTTCCTTCCA 2401
 TCCCCACCCC TACCACTAGG CCTAAGGAAT AGGGTCAGGG GCTCCAAATA GAATGTGGTT 2461
 GAGAAGTGGA ATTAAGCAGG CTAATAGAAG GCAAGGGCA AAGAAGAAAC CTTGAATGCA 2521
 TTGGGTGCTG GGTGCCTCCT TAAATAAGCA AGAAGGGTGC ATTTTGAAGA ATTGAGATAG 2581
 AAGTCTTTT GGGCTGGGTG CAGTTGCTCG TGGTTGTAAT TCCAGCACTT TGGGAGGCTG 2641
 AGGCAGGGAGG ATCACCTGAG CTTGGGAGTT CAAGACCAGC CTCACCAACG TGGAGAAACC 2701
 CTGTCTTTAC TAAAAATACA AAAAATTCACTG TGCGTCATGG TGGCACATGC CTGTAATCCC 2761
 AGCTGCTCGG GAGGCTGAGG CAGGAGAAC ACTTGAACCA GGGAGGCAGA GGTTGTGGTG 2821
 AGCAGAGATC GCGCCATGCT TCTCCAGCCT GGGCAACAAG AGCAAAAGTT CGTTAAAAAA 2881
 AAAAAAAAG TCCTTCGAT GTGACTGTCT CCTCCAAAT TTGTAGACCC TCTTAAGATC 2941
 ATGCTTTCA GATACTTCAA AGATTCCAGA AGATATGCC CGGGGGTCCT GGAAGCCACA 3001
 AGGTAAACAC AACACATCCC CCTCCTTGAC TATCAATTAA ACTAGAGGAT GTGGTGGGAA 3061
 AACCAATTATT TGATATTAAA ACAATAGGCT TGGGATGGAG TAGGATGCAA GCTCCCCAGG 3121
 AAGTTAGATA ACTGAGACTT AAAGGGTGT AAGAGTGGCA GCCTAGGGAA ATTTATCCCG 3181
 GACTCCGGGG GAGGGGGCAG AGTCACCAGC CTCTGCATTT AGGGATTCTC CGAGGAAAAG 3241
 TGTGAGAACG GCTGCAGGCA ACCCAGGCCT CCCGGCGCTA GGAGGGACGA CCCAGGCCTG 3301
 CGCGAAGAGA GGGAGAAAGT GAAGCTGGGA GTTGCCTGACT CCCAGACTTC GTTGAATGC 3361
 AGTTGGAGGG GGCAGCTGG GAGCGCGCTT GCTCCCAATC ACCGGAGAAG GAGGAGGTGG 3421
 AGGAGGAGGG CTGCTTGAGG AAGTATAAGA ATGAAGTGT GAAGCTGAGA TTCCCCCTCCA 3481
 TTGGGACCGG AGAAACCAGG GGAGCCCCCC GGGCAGCCGC GCGCCCCCTTC CCACGGGGCC 3541
 CTTTACTGCG CCGCGCGCCC GGCCCCCACC CCTCGCAGCA CCCCGCGCCC CGCGCCCTCC 3601
 CAGCCGGGTC CAGCCGGAGC CATGG

Fig. 9